DOCUMENT RESUME

ED 170 654 CG 013 445

AUTHOR Knauf, Janine; Benson, Purnell
TITLE A Peer-Rating Information System.
PUB DATE 1 Jul 78
NOTE 10p.; Paper presented at the Annual Convention of the American Psychological Association (86th, Toronto, Ontario, Canada, August, 1978)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Business Education; *Computer Oriented Programs; *Evaluation Methods; *Grade Equivalent Scores; *Grading; Graduate Students; Higher Education; *Peer Evaluation; *Rating Scales
IDENTIFIERS *PEERRATE

ABSTRACT

A peer-rating information system, designed to assess the ratings which peers make of each other's performance, is presented. An application of the computerized rating program, PEERRATE, to a graduate business course is described. Special features of the program are that: (1) rater bias is removed by considering only the intervals made by the rater between ratees; (2) these paired differences are averaged by Mosteller's procedure for combining incomplete pair comparisons; (3) an interactive procedure permits conversion of all ratings to a common scale unit; and (4) two types of interior ratings, peer and instructor, are available. (Author/HLM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

July 1, 1978

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J Knauf

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

A PEER-RATING INFORMATION SYSTEM

Janine Knauf
Purnell Benson
Rutgers University
Newark, NJ 07102

ED170654

A peer-rating information system was developed to utilize the ratings which peers make of each other's performance. This information system can be used to assess the contribution of either students or employees. It can be applied to work performed by teams or departments, as well as work by individuals. The information system is based upon the computerized rating program named PEERRATE (Benson, 1977, 1978).

In the application now reported, PEERRATE was used to grade students in a graduate business course, "Computer-Oriented Accounting and Management Information Systems" taught by Professor Knauf. The program was used for evaluating student assignments performed both individually and in groups, in the 11 sections of the course taught over a three-year period.

The computer-oriented accounting and M.I.S. course assigned each student to one of seven different areas of computerized accounting. The areas were:

(1) accounts receivable

(2) general ledger

(3) inventory

(4) billing and sales

(5) accounts payable

(6) payroll

(7) plant and equipment

CG013445

The students assigned to each of these areas were organized into teams. The students made themselves specialists in their particular areas, took part in an oral presentation by the team, and developed an individual Fortran program and information system for a hypothetical company to which the student was considered a consultant.

In addition, class members graded the teams on their oral presentations, and graded individuals on their contributions to the team effort and also on their written reports. Course grades were consequently based upon three factors.

(1) team performance, and individual contribution to this performanc

(2) individually developed accounting information systems

(3) the skill of students in rating their peers

No examinations were given in the course. Requiring students to grade the performance of their peers resulted in their close involvement in the subject matter covered by other teams. In addition, design of the grading criteria for the oral reports made these presentations both educational and enjoyable to listen to.

Rating Procedures in Detail

Each class member graded the oral presentations by teams (excepting their own team) on a scale from 50 to 99 points for each of three different categories of performance.

(1) technical proficiency in the subject matter

(2) ability to hold the class attention

(3) teaching effectiveness

These ratings along with the instructor ratings for each team for these criteria were used by the PEERRATE computer program

to generate a grade for team performance.

Then each student rated the members of his or her team, including himself or herself, for amount of contribution to the team achievement. The members of each team rated each other on a verbal-step scale.

1 = little or no contribution

2 = some contribution

3 = acceptable contribution

4 = strong contribution

5 = exceptionally strong contribution

These member-grades, together with the class rating of the team's performance, determined the grade received by each individual for the team assignment. The member-ratings were adjusted to the mean member-rating for the team. The deviation of the member-rating about this mean was then applied to the grade received by the team as a whole from the class, with whatever weight was selected by the instructor. The individual member received a grade higher than that received by the team for its overall performance if he or she was rated by other members of the team as contributing above the average, and lower, if rated below the average for the team.

This combination of team ratings and member ratings provides a rating rationale which is applicable to industry. Members of departments can be given grades which are a combination of how the department is rated by those outside the department and of how members within the department evaluate each other's performance.

The PEERRATE Computer Program

The computing procedure followed by the PEERRATE program has been reported elsewhere (Benson, 1977). Features of the program applicable to the ratings of teams or departments and the ratings made by department or team members of each other include the following:

(1) The habitual bias of the rater to be lenient or severe is removed by considering only the intervals made by the rater between ratees. That is, each rating made by the rater is subtracted from other ratings by that same rater to establish intervals between all of the pairs of ratees evaluated by that rater.

(2) These paired differences are averaged by Mosteller's procedure for combining incomplete pair comparisons (1951). If ratings by all raters are accorded equal weight, and if all pairs of ratees are rated by all raters, a simple arithmetic average would suffice to define how much above or below the group mean each ratee is scaled. Since this unique situation seldom prevails, the use of Mosteller's procedure is made routine in the PEERRATE program.

(3) Originated by Benson is an iterative procedure for converting ratings by all raters to a common scale unit. The way in which this is done is to adjust the stretch of each rater's ratings to the same stretch of ratings calculated by the program for that rater's ratees. At each iteration, the standard deviation of the rater's scale of judgment is corrected to agree with the standard deviation for that rater's ratees obtained from the PEERRATE program calculation (Benson, 1977).

(4) Also included by Benson in his program is the calculation of the average bias of each rater, the difference in stretch between the rater and a criterion, the residual error after adjusting each rater's ratings for upward or downward bias and for difference in the stretch or scale unit of the rater and a criterion. Also calculated is the absolute discrepancy between the rater's self-rating and a criterion rating.

Two kinds of criterion rating are available: that calculated from the group of peers after adjusting their ratings to a common origin and common scale unit, and that provided by the input of instructor or departmental head ratings. Calculation is made of the proficiency of the rater according to these criteria, with the different components weighted as required by the user of the PEERRATE program and converted to a 100 point scale with whatever mean and standard deviation of rater scores which the program user imposes.

The adaptation of the PEERRATE program to combined department and individual ratings was developed by Knauf, as here reported. In response to her requirements, the PEERRATE program was modified to include her provisions for rating calculations.

Results from the Grading Procedure

Grades were calculated for the students based upon the rating information supplied by both the instructor and other students. Although the PEERRATE program can function without instructor grades, all students were graded by Professor Knauf. Not only were these instructor grades weighted as half of each student's grade, but also they were used to verify the accuracy of grades as calculated by the program. The grades determined for the team presentations set the mean for each team. Then individual team members were assigned higher or lower grades depending upon how their contributions to the team were judged by their team members.

The controls for running the PEERRATE program enable the user to designate any norm and dispersion for the grades which may be required. For the grade computations reported here, the norm was set at 88 with a standard deviation of 3 on a 100 point scale. Owing to the skewness downward of the scores, their dispersion below the norm is greater than their dispersion above the norm.

Examples of the grades calculated from the inputted ratings are attached at the end of this paper.

For each of the 11 classes with which the PEERRATE program has been used to assemble grades from peer and instructor ratings, an opinion survey has been conducted to ascertain student reactions. While the proportion who favor the PEERRATE program varies from

class to class, overall about two-thirds of the students recommend continued use of the PEERRATE program, either permanently or on a further trial basis.

Extension of PEERRATE to Industrial Applications

The same methods used in this study seem equally applicable to industry. Much of the work performed in industrial settings is done by teams or task groups. Coworkers may often be better judges of each individual's contribution than the supervisor or manager. Their rating inputs can be made part of a permanent information file for the evaluation of each individual's proficiency as a worker. At the same time, the accumulated file can contain the measures of rater ability for each employee in the rating system. Periodic ratings can be inputted and reports issued for the status of each employee. The PEERRATE program can readily be used as a processing tool in a comprehensive information system which monitors each employee's performance.

REFERENCES

Benson, Purnell H. Eliminating rater biases in performance appraisal. American Statistical Association, Social Statistics Section, 1977.

Benson, Purnell H. PEERRATE: a comprehensive computerized student-grading technique. Fourth New Jersey Conference on the Use of Computers in Higher Education, Rutgers University, New Brunswick, N.J., March 20-22, 1978.

Mosteller, Frederick. Remarks on the method of paired comparisons; the least squares solution assuming equal standard deviations and equal correlations. Psychometrika, 16 (March 1951), pp. 3-9.

26A. PERF SCORES, RATING QUALITIES, AND RATER ABILITIES IN ASCENDING ORDER

PERF SCORE		RATING QUALITY		RATER ABILITY	
ID NO.	SCORE	ID NO.	SCORE	ID NO.	SCORE
1	*****	1	*****	6	*****
6	*****	6	*****	13	*****
13	*****	13	*****	14	*****
14	*****	14	*****	28	*****
28	*****	28	*****	3	82
5	80	32	84	10	84
11	80	5	85	9	84
18	80	9	85	2	85
12	80	17	85	27	85
21	82	18	85	4	85
23	82	11	85	31	85
24	82	30	85	20	86
4	82	31	85	16	86
16	84	12	85	15	87
25	84	33	85	25	87
8	84	27	86	26	87
29	84	26	86	32	87
20	88	3	89	33	87
2	88	10	89	22	88
22	88	15	89	17	88
19	88	29	89	30	89
15	91	16	89	18	89
3	91	7	89	23	90
10	91	8	89	21	90
7	91	25	89	5	90
26	95	21	90	11	91
27	95	23	90	24	91
9	95	24	90	7	91
17	95	4	90	29	91
30	96	22	93	19	93
31	96	19	93	8	93
32	96	20	93	12	93
33	96	2	93	1	98

ROE, RICHARD ID. NO. 16 PROJECT 649 NUMBER 1 ***PEER-RATING

PROF. JANINE KNAUF COMPUTER ORIENTED ACCOUNTING & MIS SUMMER 1977

**** OVERALL SCORE FOR PERFORMANCE OF ASSIGNMENT(S)---------- 84
THIS IS SCORE RECEIVED ON ASSIGNMENT WITH 55 RATERS RECORDS

POINTS FOR ITEMS ADJUSTED TO NORMS FOR GROUP:

ITEM NAME	POINT BASIS	ITEM RATING ADJ TO NORM	WEIGHT FOR ITEM
TECHNICAL CONTENT	100	82	0.167
HOLDING CLASS INTEREST	100	88	0.167
TEACHING EFFECTIVENESS	100	82	0.167
TEAM I MEMBER RATING	100	84	0.500

TEAM MEMBER SELF-RATING: 100

***** OVERALL SCORE FOR RATING PERFORMANCE---------- 86
THIS IS SCORE RECEIVED FOR RATING PERFORMANCE WITH 8 RATEES

(CALCULATIONS WHICH FOLLOW ARE DONE ON BASIS OF 100 POINTS)

SCORE FOR DEVIATION FROM ADJ GROUP RATINGS---------- 90
WEIGHT FOR DEV FROM ADJ GROUP RATINGS---------- 1.0

SCORE FOR DIFFERENCE IN RANGE FROM ADJ GROUP RATINGS------- 89
WEIGHT, DIF IN RANGE FROM ADJ GROUP RATINGS------ 1.0

SCORE FOR RESIDUAL ERROR FROM ADJ GROUP RATINGS---------- 86
WEIGHT, RESIDUAL ERROR FROM ADJ GROUP RATINGS---- 1.0

SCORE FOR ERROR IN SELF-RATINGS FROM ADJ GROUP RATINGS----------
WEIGHT, ERROR IN SELF-RTGS FROM ADJ GROUP RTGS--- 0.0
(IF BLANK, NO EVALUATION OF SELF-RATINGS IS MADE)

SCORE FOR DIFFERENCE FROM LEADER RATINGS---------- 83
WEIGHT, DEVIATION FROM LEADER RATINGS---------- 1.0

SCORE FOR DIFFERENCE IN RANGE FROM LEADER RATINGS---------- 88
WEIGHT, DIF IN RANGE FROM LEADER RATINGS--------- 1.0

SCORE FOR RESIDUAL ERROR FROM LEADER RATINGS---------- 79
WEIGHT, RESIDUAL ERROR FROM LEADER RATINGS------- 1.0

SCORE FOR ERROR IN SELF-RATINGS FROM LEADER RATINGS----------
WEIGHT, ERROR IN SELF-RTGS FROM LEADER RTGS------ 0.0
(IF BLANK, NO EVALUATION OF SELF-RATINGS IS MADE)